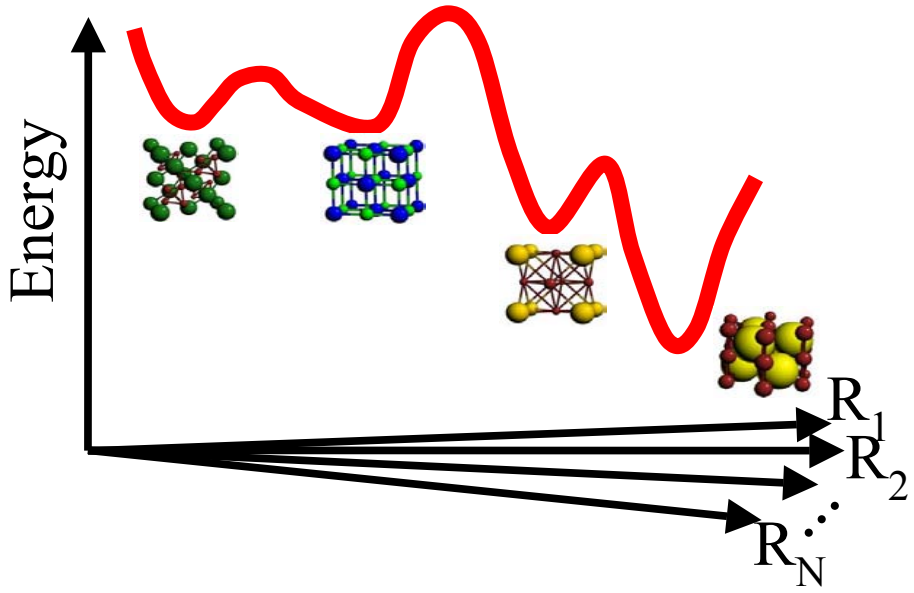
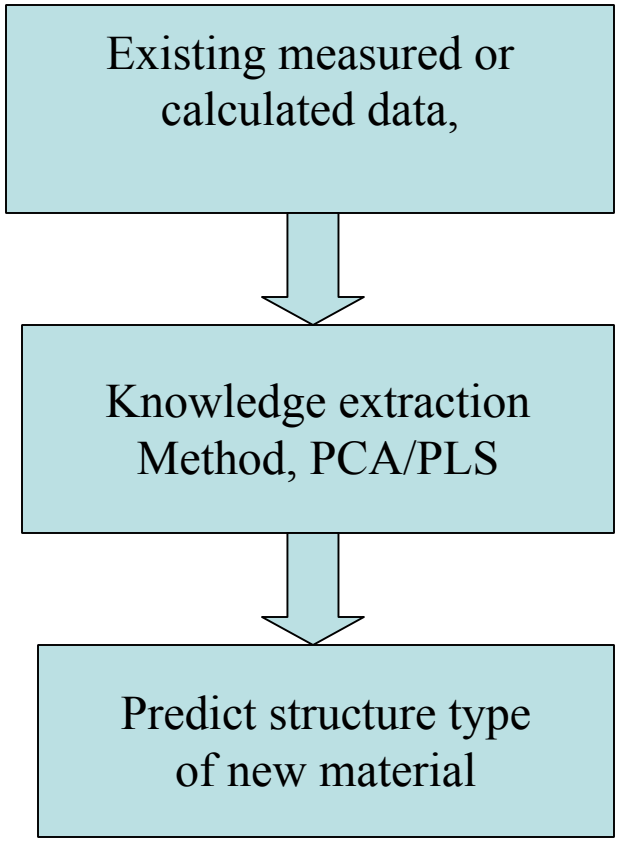
Data Mining for Predicting the Structure of Materials

Professor Gerbrand Ceder, MIT, DMR-#0312537

Predicting the stable structure type of a material is essential for understanding many of its properties. This problem, is principle intractable as there are ten of thousands of possible crystal structures. Our data mining approach uses previously known or calculated data to predict the possible stable structure types of a new material, vastly reducing the time required to idenitfy the stable structure types of a new system. Knowledge of the structure can then form the basis of a theory to predict the properties of a new material



Data Mining Approach



Data Mining for Predicting the Structure of Materials

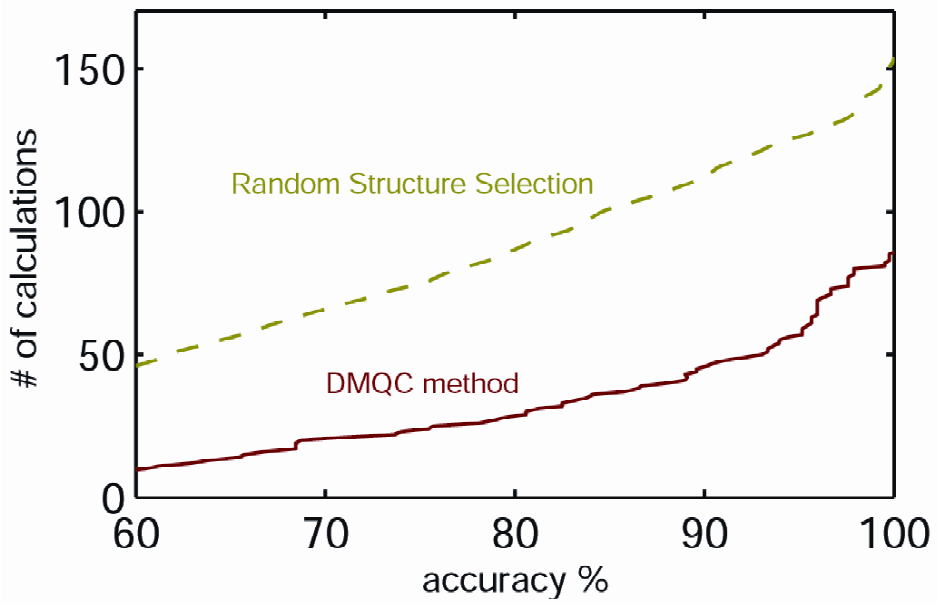
Professor Gerbrand Ceder, MIT, DMR-#0312537

Correlations among known or calculated structure type data can be used to predict the most likely stable structure types of a new material using only limited information from new data. Data Mining on Quantum Calculations (DMQC) uses a PCA with Partial Least Squares Regression on our database of 154 structure types for each of 82 binary alloys to predict likely ground state structure types

Largest database of ab-initio energies

http://datamine.mit.edu (after Oct 1, 2004)

Plot shows improvement in the ability to predict ground state structure types with DMQC over random structure selection



Broad Impact

- **The development of new materials takes a long time using traditional experimental methods**
- **Data mining methods can be used to predict a property of a new material on the basis of limited information and correlations extracted from large databases.**
- **The synergy of coupling data mining techniques to ab-initio computational models creates a powerful tool to design new materials on computer**